Mathematical Supplement: Geometry and Gravity for Weak Fields

1 Hyperbolic Functions

Hypberbolic functions relate the points on the arc of a hyperbola to the x and y coordinates defining that system. They are somewhat analogous to the trigonometric sine and cosine functions defined on a unit circle. In particular, the hyperbolic sine, sinh and hyperbolic cosine, cosh take real arguments and render real values. They are defined as

$$\sinh(a) = \frac{e^a - e^{-a}}{2} \tag{1}$$

$$\cosh(a) = \frac{e^a + e^{-a}}{2} \tag{2}$$

with the hyperpolic tangent being defined as the ratio of these two, in analogy with trigonometry.

$$\tanh(a) = \frac{\sinh(a)}{\cosh(a)} = \frac{e^a - e^{-a}}{e^a + e^{-a}} \tag{3}$$

Additional functions like coth, sech and csch can be defined in analogy with trigonometry as well.

The trigonometric sine and cosine functions are related to a circle, with their argument being the angle measured from the positive x-axis to the point on the circle in question.

$$\sin^2(\theta) + \cos^2(\theta) = 1 \tag{4}$$

The hyperbolic sine and cosine are related to a hyperbola, with their argument being twice the area in a wedge-shaped region between the positive x axis, the hyperbola, and a line from the origin to the point on the hyperbola in question; see Figure 1.

$$\cosh^2(a) - \sinh^2(a) = 1 \tag{5}$$

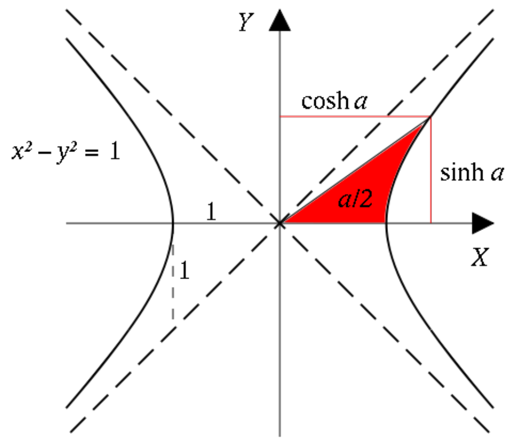

Figure 1: The hyperbolic sine and cosine take a real argument, a, twice the wedge-shaped area highlighted in red, and relate it to the $x\,(=\cosh(a))$ and $y\,(=\sinh(a))$ coordinates as shown by the red lines. Credit: Wikipedia

Graphically, the hyperbolic sine and cosine relate the area shown in red in Figure 1 to the x and y axes.

$$x = \cosh(a) \tag{6}$$
$$y = \sinh(a) \tag{7}$$

Note that the area indicated is half the argument of the sinh and cosh functions. Since they take an area as an argument, their inverse functions are sometimes referred to as the *area* hyperbolic sine, cosine, etc.

2 Matrix and Vector Operations

2.1 Matrix and Vector Multiplication

To multiply two vectors together we have

$$\boldsymbol{A} \cdot \boldsymbol{B} = A^{\mu} B_{\mu} = A^{\mu} \eta_{\mu\nu} B^{\nu} = A^0 B_0 + A^1 B_1 + A^2 B_2 + A^3 B_3$$

This is called an *inner product* of $\boldsymbol{A}$ and $\boldsymbol{B}$ and must occur between a contravariant and covariant vector, i.e., one with an upper index and one with a lower index. It is clearly symmetric and it commutes.

There is in general no cross product - it is only defined in 3-dimensions. In any case, we will not have need of an antisymmetric multiplication operator between vectors for this course.

We often have to multiply a vector by a matrix. A coordinate transformation is an example. We can use a shorthand component notation for this.

$$V^{\mu'} = A^{\mu'}{}_{\beta} V^{\beta}$$

This operation takes a vector $\mathbf{V}$ into some primed coordinate system. The vector remains the same, of course, only its coordinates are different. That is why we generally put the prime on the coordinate indices, not on the vector itself, but this is a matter of taste.

We can also write this in a different form by using bold typeface to indicate matrix and vector multiplication. In that case, the only place to put the prime is on the vector itself, but you should not be confused that $\mathbf{V}$ is different from $\mathbf{V}'$: They are the same vector, just represented in different coordinate systems.

$$\mathbf{V}' = \mathbf{A}\mathbf{V}$$

Explicitly we have

$$(V_{0'}, V_{1'}, V_{2'}, V_{3'}) = \begin{pmatrix} A_{00} & A_{10} & A_{20} & A_{30} \\ A_{01} & A_{11} & A_{21} & A_{31} \\ A_{02} & A_{12} & A_{22} & A_{32} \\ A_{03} & A_{13} & A_{23} & A_{33} \end{pmatrix} \begin{pmatrix} V_0 \\ V_1 \\ V_2 \\ V_3 \end{pmatrix} \tag{8}$$

The components of V' are

$$\begin{aligned} V_{0'} &= V_0 A_{00} + V_1 A_{10} + V_2 A_{20} + V_3 A_{30} \\ V_{1'} &= V_0 A_{01} + V_1 A_{11} + V_2 A_{21} + V_3 A_{31} \\ V_{2'} &= V_0 A_{02} + V_1 A_{12} + V_2 A_{22} + V_3 A_{32} \\ V_{3'} &= V_0 A_{03} + V_1 A_{13} + V_2 A_{23} + V_3 A_{33} \end{aligned}$$

2.2 The Trace of a Matrix

We also sometimes need to find the trace of a matrix. This is the sum of its diagonal elements. We can write this as

$$\text{trace(A)} = A^{\mu}{}_{\mu} = A^{0}{}_{0} + A^{1}{}_{1} + A^{2}{}_{2} + A^{3}{}_{3}$$

In the matrix from Equation 8, this works out to $A_{00} + A_{11} + A_{22} + A_{33}$.

The trace of a matrix is also a *contraction* of the matrix on its two indices, but we can also contract objects that have a rank larger than two. We just have to specify which indices are being contracted. So, for instance, if we had a four-dimensional third-rank tensor, $M^{\alpha}{}_{\beta\gamma}$, we could contract it on its first and third indices as follows:

$$M_{\beta} = M^{\alpha}{}_{\beta\alpha} \equiv M^{0}{}_{\beta 0} + M^{1}{}_{\beta 1} + M^{2}{}_{\beta 2} + M^{3}{}_{\beta 3} \tag{9}$$

Or if we liked, we could contract it on its first and second like this:

$$M_{\gamma} = M^{\alpha}{}_{\alpha\gamma} \equiv M^{0}{}_{0\gamma} + M^{1}{}_{1\gamma} + M^{2}{}_{2\gamma} + M^{3}{}_{3\gamma} \tag{10}$$

In each case, we end up with a tensor of rank two lower than we started with, so in these examples we end up with two vectors. These vectors are in general distinct though, as you can tell by looking at Equations 9 and 10; they have different components - unless $M^{\alpha}{}_{\beta\gamma}$ is symmetric on its second and third indices.

We cannot take a contraction of the second and third components of M unless we raise one of them; just as in the case of the dot product, contraction is only defined between an upper (contravariant) and lower (covariant) index.

2.3 Contravariant and Covariant Vectors

The terms *contravariant* and *covariant* vectors cause a lot of confusion when people begin to study general relativity. They are both just vectors in the way we generally think of them, but they are expressed in different bases, and they have different transformation properties. In contemporary work in GR "contravariant" and "covariant" have fallen out of favor. Alternate terms "dual-vector" or "one-form" have replaced covariant vector, and contravariant vectors are usually now referred to simply as vectors. The new nomenclature doesn't do anything to make the differences between the two objects less confusing, but it is less cumbersome. We will adopt it for the remainder of this section.

It is actually the need to have the contraction between two vectors on their indices yield an invariant scalar quantity that is the best way to think about the two kinds of objects. This is related to their transformation properties.

If we consider a vector, $\vec{V}$, we know that we can write it as a combination of its components and basis vectors, for some basis $\hat{a}_\beta$.

$$\vec{V} = V^\beta \hat{a}_\beta \tag{11}$$
$$= V^0 \hat{a}_0 + V^1 \hat{a}_1 + V^2 \hat{a}_2 + V^3 \hat{a}_3 \tag{12}$$

This expression gives a contravariant vector (it has a raised index), or from now on just vector, not a scalar. Each of the $\hat{a}_\beta$ are individual vectors, they are not components of vectors like the V^β. [1]

If we wished, we could write $\vec{V}$ in terms of some other basis, $\hat{a}_{\alpha'}$, where the prime on the α indicates that we are using a different set of basis vectors than before. Then we could write $\vec{V}$ in terms of this new basis.

$$\vec{V} = V^{\alpha'} \hat{a}_{\alpha'} \tag{13}$$
$$= V^{0'} \hat{a}_{0'} + V^{1'} \hat{a}_{1'} + V^{2'} \hat{a}_{2'} + V^{3'} \hat{a}_{3'} \tag{14}$$

These two expressions are the same vector, but it is represented in two different coordinate systems, each with its own basis vectors. As such, the components of the vector are different in the different coordinates.

We know that we can express either coordinate basis in terms of the other if we employ some transformation matrix, $\mathbf{\Lambda}$. For example, we could write,

$$\hat{a}_{\alpha'} = \Lambda_{\alpha'}{}^{\beta} \hat{a}_\beta \tag{15}$$
$$= \Lambda_{\alpha'}{}^{0}\, \hat{a}_0 + \Lambda_{\alpha'}{}^{1}\, \hat{a}_1 + \Lambda_{\alpha'}{}^{2}\, \hat{a}_2 + \Lambda_{\alpha'}{}^{3}\, \hat{a}_3 \tag{16}$$

If we substitute Equation 15 into Equation 13 we have the following:

$$\vec{V} = V^{\alpha'} \hat{a}_{\alpha'} = V^{\alpha'} \Lambda_{\alpha'}{}^{\beta}\, \hat{a}_\beta \tag{17}$$

Comparing Equations 17 and 11 we see that we can relate the components of $\vec{V}$ in the two coordinate bases.

$$V^\beta = V^{\alpha'} \Lambda_{\alpha'}{}^{\beta} \tag{18}$$

Comparing Equation 17 to Equation 15 we see that the transformation to go from the primed vector components to the unprimed is the same as that to go from the unprimed to the primed basis vectors. In other words, the vector components and the basis vectors have inverse transformation properties.

[1] Note: the index of a basis vector is lowered to allow the summation convention to work, and it is important so that the mathematical machinery we use will work. This does not mean that a vector basis is composed of covariant vectors (one-forms). The $\hat{a}_\beta$ are vectors.

We can perform the exact same analysis for contravariant vectors, from now on, one-forms. For instance, we can write a one-form $\tilde{P}$ in terms of basis one-forms $\hat{\omega}^\alpha$.

$$\tilde{P} = P_\beta \hat{\omega}^\beta \tag{19}$$

As before, the raised index on $\hat{\omega}^\beta$ does not mean that it is a vector. It is a basis one-form, and the raised index is simply a convenience to allow the summation convention to work with the one-form components P_β.

We could go through exactly the analysis for one-forms that we did for vectors. We would find that just as for vectors, the transformation properties of the one-form components are the inverse of the one-form bases - you can try this out if you wish. What's more, we would find that the one-forms transform inversely to the vectors. That is why they were called *contra*-variant vectors: their transformation properties were contrary to that of vectors. We have

$$V^\beta = V^{\alpha'} \Lambda_{\alpha'}{}^\beta \tag{20}$$

$$P_\beta = P_{\alpha'} \Lambda^{\alpha'}{}_\beta \tag{21}$$

And for the bases:

$$\hat{\omega}^\beta = \hat{\omega}^{\alpha'} \Lambda_{\alpha'}{}^\beta \tag{22}$$

$$\hat{a}_\beta = \hat{a}_{\alpha'} \Lambda^{\alpha'}{}_\beta \tag{23}$$

This might seem completely innocuous. It is not. If we write the contraction of $\tilde{P}$ with $\vec{V}$ we have $V^\beta P_\beta$, and we know that this is a scalar that should be invariant. So if we write it in another basis we should get the same result: $V^\beta P_\beta = V^{\beta'} P_{\beta'}$. Substituting explicitly from Equations 20 and 21 we get:

$$V^\beta P_\beta = V^{\alpha'} \Lambda_{\alpha'}{}^\beta P_{\mu'} \Lambda^{\mu'}{}_\beta \tag{24}$$

$$= V^{\alpha'} P_{\mu'} \Lambda_{\alpha'}{}^\beta \Lambda^{\mu'}{}_\beta \tag{25}$$

$$= V^{\alpha'} P_{\mu'} \delta_{\alpha'}{}^{\mu'} \tag{26}$$

$$= V^{\alpha'} P_{\alpha'} \tag{27}$$

Where $\delta_{\alpha'}{}^{\mu'}$ is the Kronecker delta function:

$$\delta_\mu{}^\nu \equiv \left\{ \begin{array}{ccc} 1 & : & \mu = \nu \\ 0 & : & \mu \neq \nu \end{array} \right. \tag{28}$$

So if we use a one-form and a vector, we get the invariance of scalar quantities that we expect. If we use two one-forms or two vectors in a similar product we will not get this invariance. You can try it by explicitly substituting as we have in Equation 24.

The most familiar one-form is probably the gradient of a scalar field. If you work out its transformation properties you will see that they are the inverse transformation properties of the position vector. This is because to take the gradient you must "divide" by the position. Try it and you will see. A (overly) general rule of thumb is, any time you divide by position you will get a one-form. If you do not, you will get a vector. The difference between the two is very often not important, but in relativity, where we cannot count on a single coordinate basis to describe all of space and time, we have to be especially careful about the transformation properties of the quantities we are working with. That is why the distinction between vectors and one-forms becomes so important when we work in GR.

3 Maxwell's Equations Are Invariant

We claimed that Maxwell's equations are invariant under Lorentz transformations. To show this we must show that $F^{\mu\nu}$ and j^{ν} transform like the four-vector position, $\mathbf{x}$, under Lorentz transformations.

Recall that in 3+1 dimensions the Maxwell equations can be written as

$$\frac{\partial F^{\mu\nu}}{\partial x^{\mu}} = \frac{4\pi}{c} j^{\nu} \tag{29}$$

and

$$\frac{\partial F^{\mu\nu}}{\partial x^{\alpha}} + \frac{\partial F^{\nu\alpha}}{\partial x^{\mu}} + \frac{\partial F^{\alpha\mu}}{\partial x^{\nu}} = 0 \tag{30}$$

The four-vector current density j^{ν} is

$$j^{\nu} \equiv (c\rho, j_x, j_y, j_z) \tag{31}$$

We will begin with the current density $\mathbf{j}$. It is the four-current and has parts made up of the normal 3-current density and the charge density, ρ.

$$\mathbf{j} \equiv (c\rho, j_x, j_y, j_z) \tag{32}$$

We can write the charge density as follows:

$$\rho = qn = qn_0\gamma \tag{33}$$

and similarly

$$j_i = qnv_i = qn_0\gamma v_i \tag{34}$$

with

$$\gamma \equiv \frac{1}{\sqrt{1-(v/c)^2}} \tag{35}$$

and q is the charge on a particle, n_0 is the number of particles per unit volume in the rest-frame of the particles and v is the speed at which the particles flow. If we have different kinds of particles with different charges we can treat each type separately, so in the interest of simplicity we will assume that all the particles are the same; it costs us no generality.

We know that the charge will satisfy the continuity equation:

$$\nabla \cdot \mathbf{J} + \frac{\partial \rho}{\partial t} = 0 \tag{36}$$

where $\mathbf{J}$ is the usual 3-dimensional current density. Charge is conserved in any given frame according to Equation 36. In other words, no charges are created or destroyed, they merely move about.

We can write the current density in terms of the four-velocity $\mathbf{U} = (\gamma, v_x\gamma, v_y\gamma, v_z\gamma)$ if we like.

$$j_\mu = qn_0U_\mu \tag{37}$$

and Equation 36 in terms of j_μ:

$$j^\mu{}_{,\mu} = 0 \tag{38}$$

Since the charge current density is the product of a scalar (rest-frame charge density) and a four-vector (four-velocity), it must itself be a four-vector. As such it is invariant under Lorentz transformations.

To show that $F^{\mu\nu}$ is invariant we must show that each of its indices transform as a four-vector. Begin by writing one of the combined Maxwell equations as it would appear in a different coordinate system:

$$\frac{\partial F^{\alpha'\beta'}}{\partial x_{\alpha'}} = \frac{4\pi}{c} j^{\beta'} \tag{39}$$

We know we can use the Lorentz transformation to write the quantities in these equations in terms of unprimed coordinates as follows:

$$\frac{\partial}{\partial x_{\alpha'}} = \Lambda^\mu_{\alpha'} \frac{\partial}{\partial x_\mu} \tag{40}$$

and

$$j^{\beta'} = \Lambda^{\beta'}_{\sigma} j^{\sigma} \tag{41}$$

We can substitute these expressions into Equation 39...

$$\Lambda^{\mu}{}_{\alpha'} \frac{\partial F^{\alpha'\beta'}}{\partial x_{\mu}} = \frac{4\pi}{c} \Lambda^{\beta'}{}_{\sigma} j^{\sigma} \tag{42}$$

and then multiply the result by $\Lambda^{\nu}{}_{\beta'}$ to get

$$\Lambda^{\nu}{}_{\beta'} \Lambda^{\mu}{}_{\alpha'} \frac{\partial F^{\alpha'\beta'}}{\partial x_{\mu}} = \frac{4\pi}{c} \Lambda^{\nu}{}_{\beta'} \Lambda^{\beta'}{}_{\sigma} j^{\sigma} \tag{43}$$

Now, using the identity $\Lambda^{\nu}{}_{\beta'} \Lambda^{\beta'}{}_{\sigma} = \eta^{\nu}{}_{\sigma}$ we arrive at

$$\Lambda^{\nu}{}_{\beta'} \Lambda^{\mu}{}_{\alpha'} \frac{\partial F^{\alpha'\beta'}}{\partial x_{\mu}} = \frac{4\pi}{c} j^{\nu} \tag{44}$$

Combining this with Equation 29 we get

$$\Lambda^{\nu}{}_{\beta'} \Lambda^{\mu}{}_{\alpha'} \frac{\partial F^{\alpha'\beta'}}{\partial x_{\mu}} = \frac{\partial F^{\mu\nu}}{\partial x_{\mu}} \tag{45}$$

Now, we know that a simple coordinate transformation must be linear, in that any change in the field in one reference frame must be mirrored in any other frame in a linear way: doubling the field in one frame must double it in another. What's more, we cannot create a field simply by changing our frame of reference: If no field exists in one frame, then no field can exist in any other. With these conditions we know that the fields in the two frames must be related in the following way

$$F^{\mu\nu} = M^{\mu\nu}{}_{\alpha'\beta'} F^{\alpha'\beta'} \tag{46}$$

for some constant coefficients $M^{\mu\nu}{}_{\alpha'\beta'}$. We thus can rewrite Equation 45 as

$$\Lambda^{\nu}{}_{\beta'} \Lambda^{\mu}{}_{\alpha'} \frac{\partial F^{\alpha'\beta'}}{\partial x_{\mu}} = \frac{\partial M^{\mu\nu}{}_{\alpha'\beta'} F^{\alpha'\beta'}}{\partial x_{\mu}} \tag{47}$$

Since the coefficients $M^{\mu\nu}{}_{\alpha'\beta'}$ are constant we can remove them from the derivative on the right hand side, leaving us with

$$\Lambda^{\nu}{}_{\beta'} \Lambda^{\mu}{}_{\alpha'} \frac{\partial F^{\alpha'\beta'}}{\partial x_{\mu}} = M^{\mu\nu}{}_{\alpha'\beta'} \frac{\partial F^{\alpha'\beta'}}{\partial x_{\mu}} \tag{48}$$

and since these equations must be true for arbitrary $F^{\mu\nu}$, it must be true that

$$\Lambda^{\nu}{}_{\beta'}\Lambda^{\mu}{}_{\alpha'} = M^{\mu\nu}{}_{\alpha'\beta'} \tag{49}$$

This last relation allows us to rewrite Equation 46 as follows...

$$F^{\mu\nu} = \Lambda^{\mu}{}_{\alpha'}\Lambda^{\nu}{}_{\beta'}F^{\alpha'\beta'} \tag{50}$$

This is exactly what we wished to show, that each of the indices of $F^{\mu\nu}$ transformed like the position four-vector. Since both $F^{\mu\nu}$ and j^{ν} are invariant under Lorentz transformations, the Maxwell's equations are as well.

4 Deriving the Riemann Tensor from Parallel Transport

By definition, the parallel transport a vector along some path, its covariant derivative along that path must vanish, or in other words,

$$\frac{\partial V^{\alpha}}{\partial x^{\beta}} + V^{\mu}\Gamma^{\alpha}{}_{\mu\beta} = 0 \tag{51}$$

We had shown that in parallel transporting a vector along a coordinate path we had a change in the vector given by

$$\delta V^{\alpha}_{x^1=a} = -\int_{x^1=a} \Gamma^{\alpha}{}_{\nu 2}V^{\nu}dx^2 \tag{52}$$

We had four similar contributions from the four path segments enclosing our region of space, as shown in Figure 2. These segments are:

$$\begin{array}{llllll}
\text{segment 1: (A to B)} & x^1 & = & a \rightarrow a+\delta a & x^2 & = & b \\
\text{segment 2: (B to C)} & x^2 & = & b \rightarrow b+\delta b & x^1 & = & a+\delta a \\
\text{segment 3: (C to D)} & x^1 & = & a+\delta a \rightarrow a & x^2 & = & b+\delta b \\
\text{segment 4: (D to A)} & x^2 & = & b+\delta b \rightarrow b & x^1 & = & a
\end{array}$$

For each segment we have a contribution like Equation 52 to the change in the vector V^{α}. Explicitly these are

$$\begin{array}{llll}
\text{segment 1: (A to B)} & \delta V^{\alpha}_1 & = & \int_{x^2=b}\Gamma^{\alpha}{}_{\nu 2}V^{\nu}dx^1 \\
\text{segment 2: (B to C)} & \delta V^{\alpha}_2 & = & \int_{x^1=a+\delta a}\Gamma^{\alpha}{}_{\nu 2}V^{\nu}dx^2 \\
\text{segment 3: (C to D)} & \delta V^{\alpha}_3 & = & -\int_{x^2=b+\delta b}\Gamma^{\alpha}{}_{\nu 2}V^{\nu}dx^1 \\
\text{segment 4: (D to A)} & \delta V^{\alpha}_4 & = & -\int_{x^1=a}\Gamma^{\alpha}{}_{\nu 2}V^{\nu}dx^2
\end{array}$$

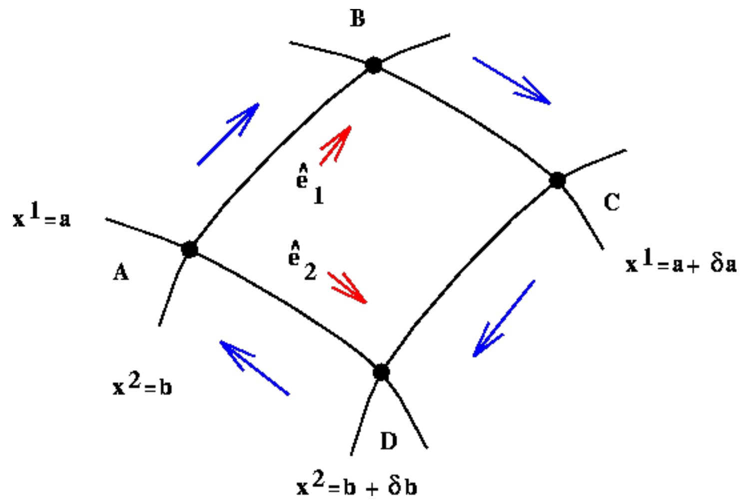

Figure 2: A possible path around which to parallel transport a vector. Credit: http://www.mth.uct.ac.za/omei/gr/chap6/node9.html.

The negative signs account for the change in direction of the path on the different segments. Adding up all these changes we will find that the difference in the vector is

$$\begin{aligned} V^{\alpha}_{final} - V^{\alpha}_{initial} = &\int_{x^2=b} \Gamma^{\alpha}{}_{\nu 2} V^{\nu} dx^1 + \int_{x^1=a+\delta a} \Gamma^{\alpha}{}_{\nu 2} V^{\nu} dx^2 \\ &- \int_{x^2=b+\delta b} \Gamma^{\alpha}{}_{\nu 2} V^{\nu} dx^1 - \int_{x^1=a} \Gamma^{\alpha}{}_{\nu 2} V^{\nu} dx^2 \end{aligned} \tag{53}$$

Since δa and δb are both small, we can expand these expressions and deal with just the first order terms. This will give us the simpler expression

$$V^{\alpha}_{final} - V^{\alpha}_{initial} = \int_a^{a+\delta a} \delta b \frac{\partial}{\partial x^2} \left[\Gamma^{\alpha}{}_{\nu 1} V^{\nu} \right] dx^1 - \int_b^{b+\delta b} \delta a \frac{\partial}{\partial x^1} \left[\Gamma^{\alpha}{}_{\nu 2} V^{\nu} \right] dx^2 \tag{54}$$

To first order we can integrate to get

$$V^{\alpha}_{final} - V^{\alpha}_{initial} = \delta b \delta a \left[\frac{\partial}{\partial x^2} \left[\Gamma^{\alpha}{}_{\nu 1} V^{\nu} \right] - \frac{\partial}{\partial x^1} \left[\Gamma^{\alpha}{}_{\nu 2} V^{\nu} \right] \right] \tag{55}$$

We can differentiate the terms inside the square brackets and replace the derivative of V^{ν} using Equation 51, relabeling the indices as necessary. After making this substitution and relabeling dummy indices we will find that the change in the vector is given by

$$V^\alpha_{final} - V^\alpha_{initial} = \delta b \delta a \left[\frac{\partial}{\partial x^2}\Gamma^\alpha{}_{\nu 1} - \frac{\partial}{\partial x^1}\Gamma^\alpha{}_{\nu 2} + \Gamma^\alpha_{\nu 2}\Gamma^\nu{}_{\mu 1} - \Gamma^\alpha_{\nu 1}\Gamma^\nu{}_{\mu 2} \right] V^\nu \tag{56}$$

This is the same equation we had before, with 1 here taking the place of σ in the main text, and 2 taking the place of λ. As before, the term inside the square brackets is the Riemann curvature tensor.

5 Gauge Transformations

Gauge transformations are changes to a function that allow it to remain a solution to some mathematical operator. They are most familiar from the study of electricity and magnetism, where we have the four Maxwell equations.

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0} \tag{57}$$

$$\nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0 \tag{58}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{59}$$

$$\nabla \times \mathbf{B} - \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} = \mu_0 \mathbf{J} \tag{60}$$

The fields $\mathbf{E}$ and $\mathbf{B}$ are related to potentials Φ and $\mathbf{A}$, respectively, as

$$\mathbf{E} = -\nabla \Phi \tag{61}$$

$$\mathbf{B} = \nabla \times \mathbf{A} \tag{62}$$

We can modify the potentials if we like as follows

$$\mathbf{A} \to \mathbf{A} + \nabla \Lambda \tag{63}$$

$$\phi \to \Phi - \frac{\partial \Lambda}{\partial t} \tag{64}$$

for an arbitrary function Λ. Any such choice of potentials will satisfy the Maxwell equations if the potentials satisfied them before the transformation was applied. For details see Jackson (1975). We are then free to choose potentials such that

$$\boldsymbol{\nabla} \cdot \mathbf{A} + \frac{\partial \Phi}{\partial t} = 0 \tag{65}$$

Any family of such solutions form a set of solutions of the Maxwell equations under the chosen gauge condition 65, which in this example is called the Lorentz gauge, though it is completely unrelated to the one for gravity having the same name. The arbitrary choice of gauge for the solutions to the linearized gravity equations is similar enough in process to the electromagnetic case that both are called *gauge transformations.*

For the gravitational case, it is not the properties of the fields we wish to understand, it is the properties of the metric tensor of the perturbed flat spacetime. In particular, we can explore the behavior of this metric under small changes in coordinate. Imagine we have

$$x^{\alpha'} = x^{\alpha} + \xi^{\alpha}(x^{\beta}) \tag{66}$$

for ξ small, of the same order as h. Then we also have

$$x^{\alpha} = x^{\alpha'} - \xi^{\alpha}(x^{\beta'}) \tag{67}$$

Differentiating these expressions we get

$$\frac{\partial x^{\alpha'}}{\partial x^{\beta}} = x^{\alpha}_{,\beta} + \xi^{\alpha}_{,\beta} \tag{68}$$

and

$$\frac{\partial x^{\alpha}}{\partial x^{\beta'}} = x^{\alpha}_{,\beta} - \xi^{\alpha}_{,\beta} \tag{69}$$

since to first order $\xi^{\alpha}_{,\beta} = \xi^{\alpha}_{,\beta'}$.

We can transform the metric using the following expression

$$g_{\alpha'\beta'}(x') = x^{\gamma}_{,\alpha'} x^{\delta}_{,\beta'} g_{\gamma\delta}(x) \tag{70}$$

Substituting Equations 68 and 69 into Equation 70 we find that we can express the metric in the new coordinates as

$$g_{\alpha'\beta'}(x') = \eta_{\alpha\beta} + h_{\alpha\beta} + \xi^{\alpha}_{,\beta} - \xi^{\beta}_{,\alpha} \tag{71}$$

This is exactly like the metric in the original coordinates, except that we must identify the perturbation as a slightly transformed perturbation $h_{\alpha'\beta'}$.

$$h_{\alpha'\beta'} = h_{\alpha\beta} + \xi^{\alpha}_{,\beta} - \xi^{\beta}_{,\alpha} \tag{72}$$

These types of transformations are called *gauge transformations* because of their similarity to the gauge transformations of electromagnetism. Note that we have four independent equations in Equations 68 and 69, so we have four parameters we can use to set a gauge.

5.1 Background Lorentz Transformations

Because we assume we are working in a nearly-flat region of spacetime, we can also perform transformations that look very much like Lorentz transformations in special relativity. Recall that Lorentz transformations, in addition to the familiar rotations or translations, can also give a boost to switch us from one inertial frame to another. So in our nearly flat space we can write

$$g_{\alpha'\beta'} = \Lambda^{\mu}{}_{\alpha'}\Lambda^{\nu}{}_{\beta'}g_{\mu\nu} = \Lambda^{\mu}{}_{\alpha'}\Lambda^{\nu}{}_{\beta'}\eta_{\mu\nu} + \Lambda^{\mu}{}_{\alpha'}\Lambda^{\nu}{}_{\beta'}h_{\mu\nu} \tag{73}$$

This equation looks suspiciously like a tensor equation, but it is not. The transformation Λ is only defined in this local region of nearly-flat spacetime, and h is definitely not a tensor. But if we confine ourselves to this Minkowski-like region of spacetime then we can treat h like a tensor. This is very convenient. For instance, if h were a tensor we could write the Riemann curvature tensor in terms of it. You can try this if you like; just substitute the nearly flat form of g into Equation 62 in Geometry and Gravity in Weak Fields (hereafter, WF). The result is to replace each g in Equation 62 (WF) with an h. Because these are not true Lorentz transformations (they are defined only on the locally flat background Minkowski spacetime, not on the entire curved spacetime of GR), we call these *background Lorentz transformations.*

6 Polarization

We wish to know how a passing gravitational wave will affect two particles on nearby geodesics. To understand this problem we can employ the equation of geodesic deviation, which tells us how adjacent geodesics[2] vary with respect to one another. In a flat spacetime like Minkowski space, geodesics that are parallel in one region will be parallel everywhere, but that is not the case when the spacetime is curved.

We can imagine two test particles are essentially at rest and separated by a displacement vector ξ^{α}. The geodesic deviation equation for these particles is

$$\nabla_U \nabla_U \xi^{\alpha} = R^{\alpha}{}_{\mu\nu\beta}U^{\mu}U^{\nu}\xi^{\beta} \tag{74}$$

The four-velocities for the particles are $U = (1, 0, 0, 0)$ to first order because they move only very slowly. The two covariant derivatives ∇_U simplify to being regular derivatives with respect to proper time because that is the only non-zero part of the four-velocities U. So we can simplify the geodesic deviation equation as follows (we write the remaining equations with the first index on R lowered, just to simplify the expressions; it makes no difference to the result)

[2] Recall that a geodesic is the spacetime path followed by an object in free fall. It is the closest thing to a straight line in curved spaces.

$$\frac{d^2\xi_\alpha}{d\tau^2} = R_{\alpha\mu\nu\beta}U^\mu U^\nu \xi^\beta \tag{75}$$

The only terms on the right hand side that do not vanish are those for $U^\lambda \neq 0$, which implies

$$\frac{d^2\xi_\alpha}{d\tau^2} = R_{\alpha 00\beta}\xi^\beta \tag{76}$$

Using Equation 62 in WF we can write the Riemann tensor in terms of the perturbations in the transverse-traceless gauge as follows

$$R_{\alpha\mu\nu\beta} = \frac{1}{2}\left(h^{TT}_{\alpha\beta,\nu\mu} + h^{TT}_{\mu\nu,\beta\alpha} - h^{TT}_{\alpha\nu,\beta\mu} - h^{TT}_{\beta\mu,\alpha\nu}\right) \tag{77}$$

$$R_{\alpha 00\beta} = \frac{1}{2}\left(h^{TT}_{\alpha\beta,00} + h^{TT}_{00,\beta\alpha} - h^{TT}_{\alpha 0,\beta,0} - h^{TT}_{\beta 0,\alpha 0}\right) \tag{78}$$

in which h^{TT} refers to the matrix in Equation 95 of WF. In the transverse-traceless gauge $h^{TT}_{\alpha 0} = 0$ for all α. This simplifies the Riemann tensor greatly.

$$R_{\alpha 00\beta} = \frac{1}{2}h^{TT}_{\alpha\beta,00} = \frac{1}{2}\frac{\partial^2}{\partial t^2}h^{TT}_{\alpha\beta} \tag{79}$$

We can now substitute this result into Equation 76, which becomes

$$\frac{\partial^2\xi^\alpha}{\partial\tau^2} = \frac{1}{2}\xi^\beta\frac{\partial^2}{\partial t^2}h^{TT\alpha}{}_\beta \tag{80}$$

Here we have raised α to match ξ^α. Let's consider the case in which the waves travel in the z direction. We know that α and β can only take on the values x or y. Furthermore, we know that $h^\alpha{}_\beta$ can take on only the values A_{11}, A_{22}, A_{12}, A_{21} from the matrix in Equation 94 in WF.

We will define $h_+ \equiv A_{11}$ and $h_\times \equiv A_{21}$ and then consider solutions of Equation 80 consistent with these definitions. There are two cases to look at: (1) $h_\times = 0$ and (2) $h_+ = 0$, and we explore these solutions in Section 6.3 in WF.

References

Jackson, J. D. 1975, Classical Electrodynamics (John Wiley $ Sons), 2nd ed.